

Pendekatan Regresi Spline Untuk Memprediksi Model Hubungan Temperatur Dengan Produksi Pythalic Anhydride Dalam Proses Reaksi Oksidasi

Nur Asiyah

Jurusan Matematika
Institut Teknologi Sepuluh Nopember, Surabaya

Abstrak

Pemodelan Regresi apabila ada asumsi bahwa bentuk fungsional dan sebaran tertentu telah diketahui dikategorikan sebagai regresi parametrik, apabila tidak ada informasi atau tidak ada asumsi-asumsi dengan model sehingga dibiarkan mengikuti perilaku data maka dikelompokkan sebagai regresi nonparametrik.

Kebutuhan aplikasi yang kompleks dengan pemodelan non linier dan kemampuan komputasi salah satunya adalah pendekatan analisis data melalui pemodelan nonparametrik yang mana menghasilkan pendekatan analisis yang mengkaji struktur data tersembunyi lebih menyeluruh dan mereduksi bias pada pemodelan parametrik. Regresi Spline dapat digunakan untuk memperoleh solusi yang optimum dari pemodelan nonparametrik. Melalui pendekatan spline ini, diaplikasikan pada proses Oksidasi, yaitu menentukan pola hubungan temperatur dan produk Phthalic Anhydride dengan katalis Vanadium Pentoksida.

Kata kunci: *Reproducing Kernel, Penalized Least Square, Spline Smothing*

1 Pendahuluan

Analisa Regresi merupakan salah satu analisis dalam statistika yang banyak digunakan untuk melihat hubungan antara variabel-variabel penjelas dan variabel respon. Diberikan data (t_i, y_i) , $i = 1, 2, \dots, n$ dan hubungan antara t_i dan y_i diasumsikan mengikuti model regresi:

$$y_i = f(t_i) + \varepsilon_i, \quad i = 1, \dots, n, \quad t_i \in [a, b] \quad (1)$$

dengan ε_i sesatan random independent $\mathcal{N}(0, \sigma^2)$, dan f fungsi regresi. Ada dua pendekatan untuk mengestimasi kurva regresi f . Bila ada informasi sebelumnya tentang bentuk kurva regresi f , berdasarkan teori atau pengalaman masa lalu, maka pendekatan regresi parametrik merupakan pendekatan yang

baik dan efisien. Tetapi sebaliknya, apabila tidak ada informasi tentang bentuk kurva regresi, maka pendekatan regresi nonparametrik dapat digunakan (Eubang, 1988). Regresi nonparametrik tidak memberikan asumsi terhadap bentuk kurva regresi. Akibatnya pendekatan ini lebih fleksibel dibandingkan dengan pendekatan regresi parametrik. Regresi spline merupakan salah satu regresi nonparametrik yang memiliki sifat tersegmen, sehingga memungkinkan untuk menyesuaikan diri secara efektif terhadap karakteristik lokal dari suatu fungsi [8]. Umumnya fungsi regresi nonparametrik diasumsikan termuat dalam suatu ruang fungsi berdimensi tak hingga. Pemilihan ruang fungsi ini biasanya dilandasi oleh sifat kemulusan (smoothness) yang diasumsikan dimiliki oleh fungsi regresi tersebut [7].

Estimator spline diperoleh dengan meminimumkan *Penalized Least Square*, yaitu kriteria estimasi yang menggabungkan *goodness of-fit* dengan *kemulusan kurva* (Ansley dkk, 1993). Oleh karena itu estimasi kurva regresi diperoleh dengan meminimumkan:

$$\min_{f \in W_2^m[a,b]} \left[\sum_{j=1}^m \{y_j - f(t_j)\} + \frac{1}{\lambda} \int_a^b \{D^m f(t)\}^2 dt \right], \lambda > 0 \quad (2)$$

dengan $W_2^m[a,b] = \{f | D^{m-1}f \text{ kontinu absolute dan } D^m f \in L_2[a,b]\}$, parameter penghalus dan D^m menyatakan operator differensial [8]. Diasumsikan setidaknya ada k knot t_i yang berbeda. Dengan pendekatan Bayesian, apabila $f(t)$ pada (1) berupa fungsi polinomial maka distribusi prior improper dibangun oleh persamaan [3].

$$f^{(m)}(t) = \sigma \lambda^{1/2} \frac{dW(t)}{dt} \quad (3)$$

Dengan $W(t)$ suatu proses Wiener dengan mean nol, dengan nilai awal terdefinisi

$$f_0 = [f(a), \dots, f^{(m-1)}(a)]^T \sim \mathcal{N}(0, kI_m) \quad (4)$$

Dalam penelitian ini dibahas mengenai permasalahan: (i) Bagaimana memperoleh estimator spline yang mempertimbangkan sifat kemulusan penalti sebagai informasi prior; (ii) Menyelidiki pengaruh Salt Bath Temperatur terhadap Produksi Phthalic Anhydride dengan metode spline.

2 Bentuk Distribusi Prior Kurva Polinomial

Pendekatan bayes untuk menghitung \hat{f} jika informasi prior berupa fungsi polinomial berderajat $m-1$

$$f(t) = b_0 + b_1 t + \dots + b_{m-1} t^{m-1}$$

maka distribusi improper prior dibangun oleh persamaan differensial stokastik (3) dengan nilai awal (4). Misalkan didefinisikan $\tilde{X}(t) = (x_1, x_2, \dots, x_n)$ dengan

$$\begin{aligned} x_1(t) &= f(t), \quad x_2(t) = x'_1(t), \quad \dots, \\ x_m(t) &= x'_{m-1}(t), \quad x'_m = \sigma \lambda^{1/2} \frac{dW(t)}{dt} \end{aligned} \quad (5)$$

maka (3) menjadi

$$\frac{d}{dt} \tilde{X}(t) = A(t) \tilde{X}(t) + \sigma \lambda^{1/2} \tilde{b} \frac{dW(t)}{dt} \quad (6)$$

(2.2) dengan $\tilde{b} = (0, \dots, 0, 1)$, $A(t)$ adalah matriks $m \times m$ dengan elemen $A_{i,i+1}(t) = 1$, $i = 1, \dots, m-1$ dan bernilai 0 untuk lainnya. Penyelesaian persamaan diferensial homogen (6) adalah

$$\phi(t) = e^{at} C$$

$$= \begin{bmatrix} 1 & t & \frac{1}{2} t^2 & \dots & \dots & \frac{1}{(m-1)!} t^{m-1} \\ 0 & 1 & t & \dots & \dots & \frac{1}{(m-2)!} t^{m-2} \\ 0 & 0 & 1 & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & t \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} C \quad (7)$$

dengan C sembarang vektor konstan. Oleh karena matriks fundamental dari (6) adalah e^{At} maka matriks transisi untuk (6) yang disebut dengan Struktur Markov (Wecker & Ansley) adalah

$$F(t, a) = e^{A(t-a)} \quad (8)$$

Persamaan (1) dan (8) dapat ditulis dalam bentuk state space [2] sebagai

$$\begin{aligned} y_i &= x_1(t_i) + \varepsilon_i \\ \tilde{X}(t) &= F(t, a) \tilde{X}(a) + \tilde{Z}(t, a) \end{aligned} \quad (9)$$

dengan $F(t, a)$ pada (8) dan $\tilde{Z}(t, a)$ adalah

$$\tilde{Z} = \int_a^t F(t, h) \sigma \lambda^{1/2} \tilde{b} dW(h) \quad (10)$$

Jika $\tilde{Z}(t, a) = (z_1, \dots, z_m)$ maka elemen ke-satu dari $\tilde{Z}(t, a)$ adalah

$$z_1(t, a) = \sigma \lambda^{1/2} \int_a^t \frac{(t-h)^{m-k}}{(m-k)!} dW(h) \quad (11)$$

Distribusi Prior dari $\tilde{X}(t)$ untuk $t \in [a, t]$ mengikuti bentuk (9) sehingga prior improper untuk $x_1(t)$, $t \in [a, t]$ (Wecker & Ansley) adalah

$$\begin{aligned} x_1(t) &= \sum_{i=0}^{m-1} \frac{(t-a)^i}{i!} \\ &+ \sigma \lambda^{1/2} \int_a^t \frac{(t-h)^{m-k}}{(m-k)!} dW(h) \end{aligned} \quad (12)$$

Tanpa mengurangi sifat keumuman dibatasi t pada interval $[0, 1]$, dan oleh karena $x_1(t) = f(t)$ dan $f(t)$ merupakan fungsi piecewise dalam interval $[0, 1]$ dengan knot t_i , $i = 1, \dots, n$ maka distribusi prior [8] diberikan oleh persamaan

$$f(t) = \sum_{i=0}^{m-1} \alpha_i \phi_i(t) + \sigma \lambda^{1/2} z_1(t) \quad (13)$$

dengan $\alpha_i = x_i(0)$, $\phi_i(t) = t^i/i!$, $i = 0, 1, \dots, m-1$, dan $z_1(t) = \int_0^1 \frac{(t-h)^{m-1}}{(m-1)!} dW(h)$

3 Estimasi Kurva Regresi Bayes

Diberikan observasi sampling $(t_1, y_1), (t_2, y_2), \dots, (t_n, y_n)$ yang diperoleh dari proses stokastik $\{Y(t) \mid t \in [a, b]\}$ pada titik-titik $0 = t_1 < t_2 < \dots < t_n = 1$. Diasumsikan fungsi regresi pada observasi ke- i mengikuti bentuk (1). Untuk proses $Y = (y_1, \dots, y_n)$ dengan $F = (f(t_1), \dots, f(t_n))$ mengikuti model

$$Y = F + \varepsilon$$

dimana $\{\varepsilon\}$ proses normal dengan mean nol dan $\text{Cov}(\varepsilon_i, \varepsilon_j) = \begin{cases} \sigma^2, & i = j, \\ 0, & i \neq j. \end{cases}$

Proses $f(t)$ diasumsikan mempunyai distribusi prior persamaan (13). Sedangkan $\alpha = (\alpha_0, \dots, \alpha_{m-1})^T$ berdistribusi normal dengan mean nol dan varian covarian kI . Misalkan $Q_1 = \sigma \lambda^{1/2} z_1(t, a)$ proses normal (Wecker, Ansley, 1983) dengan mean nol dan covarian $R^1(s, v)$ dengan

$$R^1(s, v) = \sigma^2 \lambda \int_0^1 \frac{(s-u)_+^{m-1} (v-u)_+^{m-1}}{\{(m-1)!\}^2} du \quad (14)$$

Dari persamaan (9) diperoleh $X(t) = F = T^T \alpha + \sigma \lambda^{1/2} Z(t)$, dengan $T_{m \times n}$ adalah matriks $[\phi_i(t_j)]_{i=1, m}^{j=1, n}$, selanjutnya didapat $E(FF^T)$ yaitu

$$\begin{aligned} E(FF^T) &= E\left(T^T \alpha + \sigma \lambda^{1/2} Z(t)\right) \left(T^T \alpha + \sigma \lambda^{1/2} Z(t)\right)^T \\ &= \sigma \lambda^{1/2} [(\sigma^2 \lambda)^{-1} T^T k I T + \sigma^2 \lambda R^1(t)] \\ &= \sigma^2 \lambda R_{FF}^1 \end{aligned} \quad (15)$$

Sedangkan Covarian error

$$\text{Cov}(\varepsilon) = E(\varepsilon \varepsilon^T) - (E(\varepsilon))^2 = \sigma^2 I$$

Lemma 3.1 Diketahui Y, F, ε vektor random berdistribusi n variat yang mempunyai hubungan $Y = F + \varepsilon$, dengan $E(F) = E(\varepsilon) = E(F \varepsilon^T) = 0$, $E(FF^T) = \sigma^2 \lambda R_{FF}^1$, $E(\varepsilon \varepsilon^T) = \sigma^2 I$. Jika H berdistribusi n variat normal dan memenuhi $E(H) = E(H \varepsilon^T) = 0$, $E(HF^T) = \sigma^2 \lambda R_{FF}^1$, $E(HH^T) = \sigma^2 \lambda R_{HH}^1$, maka

$$E[H|Y] = R_{HF}^1 (R_F^1 + \frac{1}{\lambda} I)^{-1} Y \quad (16)$$

Bukti: Distribusi multivariat normal bersyarat dari H jika diberikan Y , adalah

$$E[H|Y] = E(H) + \Sigma_{12} \Sigma_{22}^{-1} (Y - E(Y)) \quad (17)$$

Dengan

$$\begin{aligned} \Sigma_{12} &= E(H - E(H))(Y - E(Y))^T \\ &= \sigma^2 \lambda R_{HF}^1 \end{aligned}$$

$$\begin{aligned} \Sigma_{12} &= E(Y - E(Y))(Y - E(Y))^T \\ &= E(FF^T) + E(\varepsilon \varepsilon^T) \\ &= \sigma^2 \lambda R_{FF}^1 + \sigma^2 \\ &= \sigma^2 \lambda (R_F^1 + \frac{1}{\lambda} I) \end{aligned}$$

maka

$$\begin{aligned} E[H|Y] &= \sigma^2 \lambda R_{HF}^1 [\sigma^2 \lambda (R_F^1 + \frac{1}{\lambda} I)]^{-1} Y \\ &= \sigma^2 \lambda R_{HF}^1 (\sigma^2 \lambda)^{-1} (R_{HF}^1 + \frac{1}{\lambda} I)^{-1} Y \\ &= R_{HF}^1 (R_F^1 + \frac{1}{\lambda} I)^{-1} Y \quad \square \end{aligned}$$

Teorema 3.2 Mean posterior dari $f(t)$ adalah

$$\begin{aligned} E(f(t)|Y) &= \phi^T K T (K T^T T + M)^{-1} Y \\ &\quad + \psi^T (K T^T T + M)^{-1} Y \end{aligned} \quad (18)$$

dengan $T_{m \times n}$ adalah matriks $[\phi_i(t_j)]_{i=1, m}^{j=1, n}$, $M_{n \times n} = \frac{1}{\lambda} I + \Sigma$, $\Sigma = [R^1(t_j, t_j)]_{i=1, m}^{j=1, n}$, $R^1(s, v) = \int_0^1 \frac{(s-u)_+^{m-1} (v-u)_+^{m-1}}{\{(m-1)!\}^2} du$, $\psi = (R^1(t, t_i))_{i=1, n}^T$, dan $K = (\sigma^2 \lambda)^{-1} k$

Bukti: Didefinisikan

$F = (f(t_1), \dots, f(t_n))^T$ dan $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$ mbox

Apabila diambil $H = f(t)$ dan dari (15) diperoleh

$$\begin{aligned} R_{FF}^1 &= (\sigma^2 \lambda)^{-1} E(FF^T) \\ &= (\sigma^2 \lambda)^{-1} E[(T^T \alpha + \sigma \lambda^{1/2} Z)(T^T \alpha + \sigma \lambda^{1/2} Z)^T] \\ &= (\sigma^2 \lambda)^{-1} k T^T T + R^1 = K T^T T + \Sigma \end{aligned}$$

$$R_{f(t)F}^1 = (\sigma^2 \lambda)^{-1} E(f(t) F^T)$$

$$= (\sigma^2 \lambda)^{-1} E[(\phi^T \alpha + \sigma \lambda^{1/2} Z_1(t))(T^T \alpha + \sigma \lambda^{1/2} Z)^T] \lim_{K \rightarrow \infty} KT(KT^T + M)^{-1}$$

$$= (\sigma^2 \lambda)^{-1} k \phi^T T + (R^1(t, t_1), \dots, R^1(t, t_n))$$

$$= K \phi^T T + \psi^T$$

Berdasarkan Lemma 3.1

$$\begin{aligned} E[f(t)|Y] &= R_{f(t)F}^1 \left(R_F^1 + \frac{1}{\lambda} I \right)^{-1} Y \\ &= \phi^T KT((KT^T T + M)^{-1} Y \\ &\quad + \psi^T (KT^T T + M)^{-1} Y. \quad \square \end{aligned}$$

Untuk memperoleh fungsi estimasi \hat{f}_λ diberikan teorema berikut (Wahba, G, 1978).

Teorema 3.3 Jika $f(t)$, $t \in [0, 1]$ mempunyai distribusi prior (13) dengan $\tilde{\alpha} \sim \mathcal{N}(0, kI_m)$, $\phi_j(t) = t^j/j!$, $j = 0, 1, \dots, m-1$ dan $Z_1(t)$ proses Weiner dengan $z_1(t) = \int_a^t \frac{(t-h)^{m-1}}{(m-1)!} dW(h)$, maka spline polinomial $f_\lambda(t)$ yang meminimumkan

$$\sum_{i=1}^n (y_i - f(t_i))^2 + \frac{1}{\lambda} \int_0^1 (f^{(m)}(t))^2 dt$$

adalah

$$\begin{aligned} J_\lambda(t) &= \lim_{k \rightarrow \infty} E_k(f(t) | Y) \\ &= \phi^T (TM^{-1} T^T)^{-1} TM^{-1} Y \\ &\quad + \psi^T M^{-1} [I - T^T (TM^{-1} T^T)^{-1} TM^{-1}] Y \end{aligned}$$

Bukti: Pandang kembali (18)

$$\begin{aligned} E(f(t)|Y) &= \phi^T KT(KT^T T + M)^{-1} Y \\ &\quad + \psi^T (KT^T T + M)^{-1} Y \end{aligned}$$

Selanjutnya diperoleh (Schott, 1997):

$$\begin{aligned} &(KT^T T + M)^{-1} \\ &= M^{-1} - M^{-1} KT^T (TMT^T)^{-1} [(TMT^T)^{-1} \\ &\quad + TM^{-1} KT^T (TMT^T)^{-1}]^{-1} TM^{-1} \\ &= M^{-1} [I - T^T (TM^{-1} T^T)^{-1} \\ &\quad \cdot [(KTMT^{-1} T^T)^{-1} + I]^{-1} TM^{-1}] \end{aligned}$$

dan akibatnya

$$\begin{aligned} &\lim_{K \rightarrow \infty} (KT^T T + M)^{-1} \\ &= \lim_{K \rightarrow \infty} M^{-1} \{ I - T^T (TM^{-1} T^T) TM^{-1} \\ &\quad + T^T K^{-1} (TM^{-1} T^T)^{-2} TM^{-1} \\ &\quad - T^T K^{-2} (TM^{-1} T^T)^{-3} TM^{-1} + \dots \} \\ &= M^{-1} \{ I - T^T (TM^{-1} T^T)^{-1} TM^{-1} \} \end{aligned}$$

Demikian pula suku pertama dari ruas kanan diperoleh

$$\begin{aligned} &= \lim_{K \rightarrow \infty} [KTMT^{-1} \{ I - T^T (TM^{-1} T^T)^{-1} TM^{-1} \\ &\quad + T^T K^{-1} (TM^{-1} T^T)^{-2} TM^{-1} \\ &\quad - T^T K^{-2} (TM^{-1} T^T)^{-3} TM^{-1} + \dots \}] \\ &= \lim_{K \rightarrow \infty} [KTMT^{-1} - KTMT^{-1} \\ &\quad + (TM^{-1} T^T) TM^{-1} \\ &\quad - K^{-1} (TM^{-1} T^T)^{-2} + \dots] \\ &= (TM^{-1} T^T)^{-1} TM^{-1} \end{aligned}$$

akibatnya

$$\begin{aligned} (f(t)|Y) &= \phi^T (TM^{-1} T^T)^{-1} TM^{-1} Y \\ &\quad + \psi^T M^{-1} \{ I - T^T (TM^{-1} T^T)^{-1} TM^{-1} \} Y \quad \square \end{aligned}$$

4 Aplikasi Regresi Non-parametrik Spline

Bentuk estimator spline (14) merupakan kurva polinomial kontinu sepotong-sepotong (*piece-wise*) dengan beberapa batasan kontinuitas. Spline order m dengan knots ζ_1, \dots, ζ_k adalah suatu fungsi yang mempunyai bentuk:

$$f(t) = \sum_{i=0}^{m-1} \alpha_i t^i + \sum_{i=1}^k \delta_i (t - \zeta_i)_+^{m-1} \quad (19)$$

dengan $\alpha_0, \dots, \alpha_{m-1}$, $\delta_1, \dots, \delta_k$ koefisien-koefisien bernilai real. Spline order m mempunyai sifat (Eubank, 1988):

- f merupakan polinomial piecewise derajat $m-1$ pada tiap subinterval (ζ_i, ζ_{i+1}) .
- f mempunyai turunan yang kontinu tingkat $m-2$.
- f memiliki turunan ke $m-1$ berupa fungsi step dengan lompatan pada ζ_1, \dots, ζ_k .

Pemilihan nilai λ optimal dari (2) akan didapatkan estimator spline yang optimal yang mana sangat dipengaruhi oleh lokasi titik-titik knots (Suatu titik dimana terjadi perubahan gradien yang sangat besar). Pemilihan knots untuk mendapatkan nilai λ optimal diperoleh dari minimum $GCV(\lambda)$ yang didefinisikan:

$$GCV(\lambda) = \sum_{k=1}^n \frac{(y_k - L_k f_\lambda)^2}{(n - \text{Tr} A(\lambda))^2}$$

Jika λ optimal adalah nilai λ dengan titik-titik knots ζ_1, \dots, ζ_k , yaitu:

$$\lambda_{\text{opt}} = (\zeta_1, \dots, \zeta_k)$$

maka estimator spline f_λ dapat diperoleh dengan menggunakan metode Kuadrat terkecil. Didefinisikan: $x_1(t) = 1, \dots, x_{m+1}(t) = (t - \zeta_1)_+^{m-1}, \dots, x_{m+k}(t) = (t - \zeta_k)_+^{m-1}$. Penduga kuadrat terkecil dari f adalah:

$$f_\lambda = \sum_{j=1}^{m+k} \beta_{\lambda_j} x_j \quad (20)$$

dengan $\beta_\lambda = (\alpha_{\lambda_0}, \dots, \alpha_{\lambda_{(m-1)}}, \delta_{\lambda_1}, \dots, \delta_{\lambda_k})$ yang diperoleh dari meminimumkan

$$\text{MSE}(\beta, \lambda) = \sum_{i=1}^n \left[y_i - \sum_{j=1}^{m+k} \beta_{\lambda_j} x_j(t_i) \right]^2 \quad (21)$$

Jika $X(\lambda) = \{x_j(t_i)\}_{j=1, m+k}^{i=1, n}$, dan merupakan matriks dengan rank penuh maka β_λ adalah penyelesaian persamaan normal:

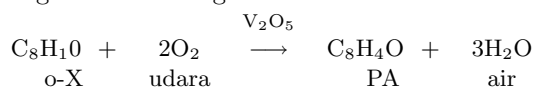
$$X(\lambda)^T X(\lambda) \beta = X(\lambda)^T Y$$

Persamaan normal ini memberikan:

$$\beta_\lambda = [X(\lambda)^T X(\lambda)]^{-1} X(\lambda)^T Y, \quad (22)$$

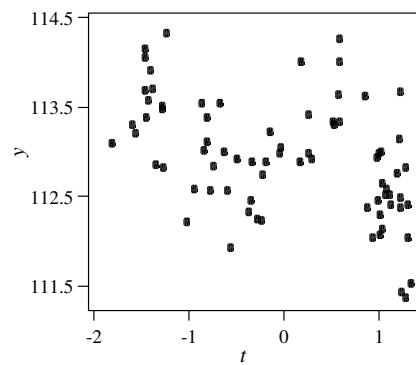
dengan $Y = (y_1, \dots, y_n)^T$.

Berikut diberikan suatu aplikasi dari model spline pada proses oksidasi reaksi kimia antara Ortho-Xylene dengan udara yang menghasilkan Phthalic Anhydride (PA) [4]. Uraian singkat Proses produksi PA sebagai berikut. Ortho-Xylene diterima terus menerus dari tangki o-X dan diumpungkan ke campuran Ortho-Xylene dan udara dengan menggunakan pompa aliran Ortho-Xylene dan dikendalikan rasionya dengan aliran udara. Ortho-Xylene sebelum masuk pada mixer dipanaskan 130°C dalam preheater sedangkan udara atmosfer dihisap oleh blower dan ditekan sampai 170°C , dimana aliran udara dikontrol dengan variasi Turbine Speed. Reaksi terjadi di Reaktor yang dilengkapi dengan tube-tube dimana tube diisi dengan katalis Vanadium Pentoksida (V_2O_5) dengan reaksi sebagai berikut



Pada kondisi normal reaksi dengan katalis Vanadium Pentoksida akan menghasilkan PA

minimum 112% dari 100% o-X yang direaksikan. Supaya hasil reaksi sesuai dengan spesifikasi maka temperatur reaksi oksidasi SBT (*Salt Bath Temperature*) harus dijaga pada kondisi optimum. Plot antara temperatur (t) dan Produksi PA (y) yang telah distandarkan disajikan dalam Gambar 1. Berdasarkan plot Gambar 1, terlihat bahwa tidak ada pola hubungan yang jelas antara temperatur dengan produksi PA. Terdapat kecenderungan pola hubungan tersebut berubah antara temperatur yang kurang dari -1.3 , temperatur antara -1.3 sampai 0.2 dan temperatur yang lebih dari 0.2 . Berdasarkan pada realita ini digunakan pendekatan spline untuk model hubungan antara temperatur dan produksi PA. Untuk mendapatkan titik-titik knots optimal digunakan metode *GCV*. Ringkasan nilai *GCV* untuk berbagai titik knots diberikan dalam Tabel 1.



Gambar 1

Tabel 1: Ringkasan nilai *GCV* untuk berbagai titik knots

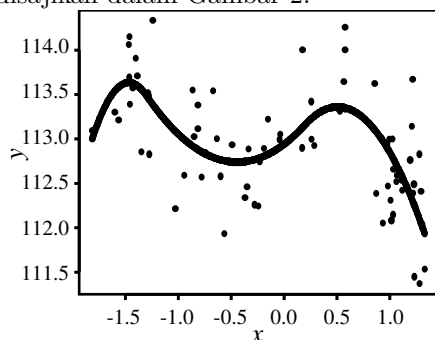
TITIK KNOTS	NILAI <i>GCV</i>
-1.20, 1.20	0.3438399
-1.20, 0.20	0.2593792
-1.30, 0.20	0.2589715
-1.40, 0.20	0.2597656
-1.50, 0.20	0.2640690
-1.53, 0.20	0.2654958
-1.53, 0.50	0.2828348
-1.53, 0.60	0.2962743
-1.53, 0.70	0.3114871
-1.53, 0.80	0.3253901

Terlihat dari Tabel 1, nilai *GCV* yang terkecil diberikan oleh titik knots -1.3 dan 0.2 , dengan nilai *GCV* = 0.2589715. Berdasarkan titik-titik knots ini didapat model spline kuadrat

dengan dua knots dapat ditulis sebagai

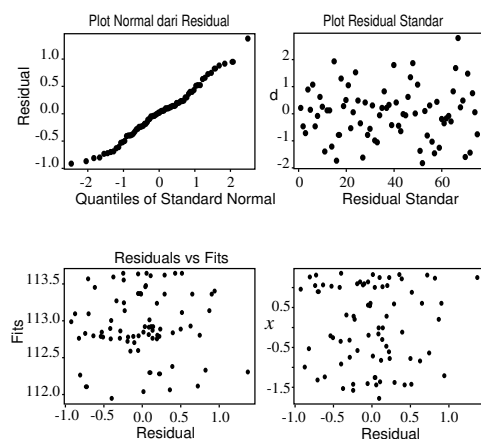
$$\hat{Y} = \beta_0 + \beta_1 z + \beta_2 z^2 + \beta_3 (z + 1.3)_+^2 + \beta_4 (z - 0.2)_+^2$$

Berdasarkan metode kuadrat terkecil didapatkan model estimasi spline: $\hat{Y} = 102.2519 - 15.53653z - 5.303263z^2 - 6.321492(z + 1.3)_+^2 - 3.131489(z - 0.2)_+^2$. Plot data dan estimasi kurva spline proses oksidasi dengan dua knots disajikan dalam Gambar 2.



Gambar 2

Selanjutnya diselidiki signifikansi dari model spline kuadrat dengan dua titik knots. Tabel 2 dan Tabel 3 (dalam Lampiran) berturut turut menyajikan Ringkasan ANOVA dan estimasi koefisien-koefisien dari model spline kuadrat.



Gambar 3

Pertama dilakukan uji hipotesis,

$$H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

$$H_1 : \text{Tidak semua } \beta_j = 0, j = 0, 1, 2, 3, 4.$$

Dengan tingkat Signifikansi $\alpha = 5\%$ dapat diambil kesimpulan bahwa hipotesis H_0 ditolak. Jadi tidak semua $\beta_j = 0, j = 0, 1, 2, 3, 4$. Selanjutnya dilakukan uji Hipotesis individu:

$$H_0 : \beta_0 = 0, \quad H_1 : \beta_0 \neq 0$$

$$H_0 : \beta_1 = 0, \quad H_1 : \beta_1 \neq 0$$

$$\vdots$$

$$H_0 : \beta_4 = 0, \quad H_1 : \beta_4 \neq 0$$

Dengan tingkat signifikansi 5% dapat diambil kesimpulan semua j tidak sama dengan nol, $j = 0, 1, 2, 3, 4$.

Langkah berikutnya adalah melakukan analisis residual untuk model spline kuadrat dengan dua titik knots. Analisis residual untuk model spline diberikan dalam Gambar 3. Terlihat dari Gambar 3, bahwa tidak ada indikasi penyimpangan dari model spline kuadrat. Dengan demikian model pendekatan nonparametrik spline kuadrat dengan dua titik knots cukup memadai sebagai salah satu model pendekatan untuk menyelidiki pola hubungan antara temperatur dan produksi PA.

5 Kesimpulan

- Estimasi kurva regresi dengan metode spline diberikan oleh Teorema 3.3.
- Model pola hubungan antara temperatur dengan produksi Phythalic Anhidride mengikuti bentuk:

$$\hat{Y} = 102.2519 - 15.53653z - 5.303263z^2 - 6.321492(z + 1.3)_+^2 - 3.131489(z - 0.2)_+^2$$
 Uji Hipotesis maupun uji individu signifikan dan tidak ada indikasi penyimpangan model.

Daftar Pustaka

- [1] Anik, *Analisis Optimasi Salt Bath Temperature (SBT) dan Mix Ratio dalam proses reaksi Gas pada produksi Phythalic Anhidride dan Maleic Anhidride di PT. Petro Widada*, Tugas Akhir Sarjana, Jurusan Statistika, ITS, 1999.
- [2] Ansley, C.F. & Khon, R., "Estimation, Filtering and smothing in state space models with incompletely specified initial condition," *The Annal. of Statistics*, 13, (1985) 1286–1316.
- [3] Ansley, C.F, Kohn, R., and Wong, C.M, "Nonparametric spline Regression with Prior Information," *Biometrika*, 80, 1, (1993) 75–88.

- [4] Asiyah, N., *Regresi Non Parametrik Spline dengan Informasi Prior Kurva Plonomial*, Tesis, ITS, 2002.
- [5] Budiantara, I.N. & Subanar, *Regresi Spline dan Permasalahannya*, Naskah Publikasi, UGM, 1996.
- [6] Craven, P., Wahba, G., "Smoothing Noisy Data with Spline Functions," *Numerrische Mathematic*, 31, (1979) 377–403.
- [7] Kimeldorf, G.S & Wahba, G., "Some Result on Tchebycheffian spline", *The Annals. of Mathematic*, 33, (1971) 82–95.
- [8] Wahba, G., *Spline Models for Observational Data*, Society for Industrial and Aplied Mathematics, Philadelphia, Pensilvania, 1990.

Lampiran

Tabel 2: ANOVA model spline kuadrat

SV	df	SS	MS	F-Hit	F-Tabel
Regresi	5	14.285	2.856999	11.82011	2.345586
Galat	70	16.91947	0.2417067		
Total	75	31.20447			

Tabel 3: Estimasi koefien-koefisien model spline

Koefisien	Estimasi	Stdev	t-hit
β_0	102.2519	0	\sim
β_1	-15.53653	0	\sim
β_2	-5.303263	1.117595	-4.745247
β_3	6.321492	1.171767	5.394838
β_4	-3.131489	0.274033	-11.42742

nilai tabel t 95%: 1.994437 dan Koef. Determinasi: 0.4577869